Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at http://about.jstor.org/participate-jstor/individuals/early-journal-content.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

MATHEMATICS AND PSYCHOLOGY.

By C. C. Grove.

(Continued from page 10, Vol. IX.)

Part II.

The Mathematics of Psychologists and the Validity of their Uses of its Forms and Processes.

The "measurements" herein spoken of as ***indirect*** are so called because they are physical measurements upon the physical causes or effects of mental activity. "They derive all their significance from the correlated psychical processes," says Brown, l. c., p. 11. They belong to the period of investigation that may be referred to as that of "mental tests." Their name is legion. This came about owing to the early misconception that any psychological tests would serve the purpose of "mental tests," and also to the fact that many without a clear or an adequate conception of just what they were testing went on applying their inventive genius to devising test after test as if to outdo all other workers in the cleverness of their tests. Prominent workers even presupposed that this or that was the essential element and measure of a particular mental process, without testing whether there was any correlation even between them. Reference is made to this in the excellent monograph "The Psychological Methods of Testing Intelligence," by William Stern, translated by G. M. Whipple. Warwick & York, Inc., Baltimore, 1914, in the words (p. 24): Very seldom was any actual preliminary investigation made to see whether this particular test was really superior to so and so many others by virtue of the precision, constancy and significance of the particular values that it afforded.

The day has not yet passed of devising single tests or series of tests without apparent concern as to what is really being tested or as to the scientific superiority of the test, for only last

January across there at the University of Pennsylvania we heard Section L of the A. A. A. S. enthusiastically introduced to a test to have been published last September, whose chief charm seemed to rest in the fact that "the children fairly ate it up."

William Stern laid a scientifically firm foundation by first defining clearly what he was testing, viz., Intelligence, l. c., page 3:

Intelligence is a general capacity of an individual consciously to adjust his thinking to new requirements: it is general mental adaptability to new problems and conditions of life.

By this definition he is enabled to distinguish intelligence from *memory,* which has to do not with the *new* but with the past, from *genius,* which is characterized not by *adaptability* but by creation, and finally from *talent,* which is not *general* but specific *capacity.*

The methods of testing intelligence are divided in the monograph into four types—*Single Tests, Series of Tests,* the *Method of Age Gradation,* and the *Method of Rank Correlation.*

On pages 13–18, he states and discusses four main classes of *single tests,* both indirect and direct, and expresses his conviction of the inadequacy of the single test by "laying emphasis upon a general critical position: *no single test, no matter how good it may be, should ever be made the instrument for testing the intelligence of an individual.*"

All this, the definition, discussion and conclusions, appeals very strongly to the present writer because of its clarity and truly scientific spirit. Several quotations are now inserted because they bring corroboration from an eminent psychologist to the writer's own views to follow.

P. 8: The consequent demand for the creation of normal test-series for each year of childhood was met, as a matter of fact, not from the side of psychiatry, but from that of psychology. Alfred Binet, with the co-operation of the physician, Simon, has created such a graded series of tests; and although the system as it now stands may be far from final, its fundamental conception will retain its permanent value and will doubtless lead us ultimately to a completely satisfactory solution.

P. 26: This demand for a system of tests presents such an exceedingly difficult scientific problem that it is perfectly evident that alienists and educators can not solve it as a side issue of their professional work, but that psychology itself will have to undertake the task. It is interesting in this connection to note how psychology attacked the problem along two very different lines: The *method of age gradation of Binet and Simon permits of a rough* gradation of intelligence for the whole range of development of the child; it *is for use in a comparable manner* with children of different ages, of different nationality and cultural level, with normal and with feebleminded children of all grades. The *method of rank correlation* is limited thus far to a comparison of the members of a small homogeneous group, but renders it possible to test the gradation of intelligence within this group with a precision that the Binet method cannot approximate.

Pp. 11, 12: But just this anticipated extension of the practical applicability of intelligence tests necessitates several words of warning.

(*a*) We are still in the midst of our *preliminary work* on method. The methods that now prevail—and this is true also of the Binet-Simon system—are not yet to be regarded as diagnostic canons that admit of official prescription. The law passed in New Jersey that directs the use of intelligence tests with all pupils suspected of backwardness seems on this account very premature. So, too, it will be long, very long, before we realize the optimistic hope that Spearman attaches to the correlation method of testing intelligence, when he says: "Indeed, it seems possible to forsee the day when there will be an annual official determination of the 'intellectual index' of every child in the empire."

(*b*) It must be understood that tests of intelligence are not easy to conduct. Their administration demands extended practice, psychological training, and a critical mind. Thus, for instance, the average teacher, whose work has been with the wholly different methods of pedagogical questioning and examining, is very apt to apply psychological tests in those forms in which their value would be positively illusory. If, accordingly, the use of tests for practical purposes shall attain any very large currency, the training of a specially psychologically drilled personnel will become a necessity. School psychologists would then take their place side by side with the school physician.

(*c*) Psychological tests ***must not be overestimated,*** as if they were complete and automatically operative measures of mind. At most they are the psychographic minimum that gives us a first orientation concerning individuals about whom nothing else is known, and they are of service to complement and to render comparable and objectively gradable other observations —psychological, pedagogical, medical—not to replace these. (See *Jour. of Educ. Psychol.,* 2: 1911, 121 and 191 by J. E. W. Wallin; also "Manual of Mental and Physical Tests," by G. M. Whipple.)

P. 33: The individual tests are of unequal value.

P. 34: The ***technique*** of the Binet-Simon method is by no means so easy as the simplicity of the material used would lead one at first to suppose. It is to be recommended that, so far as is in any way feasible, the examiner should always do his work with the aid of an assistant to keep the record, so as to avoid the undesirable division of attention between testing and recording. Both ***these experimenters must have gained a high degree of practice*** and be well used to one another ***before they proceed to actual testing.*** The examiner must have an almost mechanical exactness and uniformity in the formulation of the continually recurring questions, in the modulation of his voice, etc., yet he must be prepared for the many individual variations that appear in consequence of different reactions of the subjects, and must have definite measures in readiness for use in these junctures. Never must he permit it to be seen that some answers are more, others less satisfactory to him: rather must he maintain an attitude of uniform and quiet friendliness. The recorder should not confine himself to the mere recording of plus and minus signs to show the net outcome of each test, but should also note down as fully as possible what the subject says and also such features of his behavior and attitude towards the tests as are worth noting. This is necessary both because it is often impossible to decide whether to credit a test "plus" or "minus" until later on, after quiet consideration (and the material must be available for that) and also because it should make possible a qualitative analysis of the examinee.

P. 40: The value of observing the child during the testing must not be underestimated, for in many of the tests there are ways of setting about the task that may be of great interest (and for medical and pedagogical judgment of the case, too), though these things would not be evident from the mere plus or minus set down for the outcome of the tests.

P. 49: The *international accordance* in the judgment as to special ease or special difficulty of certain test levels, despite the differences in race and language, despite the divergences in school organization and in methods of instruction, is, in my opinion, the best vindication of the *principle* of the tests that one could imagine, because this agreement demonstrates that *the tests do actually reach and discover the general developmental conditions of intelligence* (so far as these are operative in public school children of the present cultural epoch), and not mere fragments of knowledge and attainments acquired by chance.

P. 58: Tests of intelligence have already taught us that the relations between intelligence and school ability are by no means so strict and uniform as most persons had thought.

P. 63: Complete agreement between school ability and intellectual ability is not to be expected at all nor even to be desired, *because performance in the school depends not only upon intelligence, but also upon certain other and quite different factors.* . . . But besides this there are concerned factors that have nothing at all to do with intellect, but belong to the domain of will, in the widest sense of that term: I mean the degree and duration of attention, industry and conscientiousness, sense of duty and capacity to fit into the social group.

P. 64: Intelligence is never more than a partial factor in school activity: and this demonstration may serve to refute that one-sided intellectualism that notes and values in pupils only their intellectual ability. Not that intellectual endowment is not still to be regarded as a factor of chief importance: in truth when by tests of intelligence and other psychological devices we shall have obtained a more exact knowledge of it, there will be much of profit for the schools and many mistakes and wrong courses of procedure can be prevented, and this so much the more as we get clear ideas of the range and limits of its meaning and importance. If, for instance, a given pupil shows only a moderate success in the tests of intelligence but does distinctly good work at school, and if there is no chance that a special talent might have exerted a decided influence (which could easily be recognized if existent), then there is a probability approximating to certainty that this pupil's strength is to be sought primarily in qualities of character and will.

P. 9: . . . we must guard against the danger . . . of supposing that we have grasped the individuality of a pupil in its totality when we have tested his intelligence. The fact that intelligence can be more easily treated quantitatively than can other individual capacities must not lead us to overestimate its import. Nevertheless, the fact that we can deal with intelligence by itself does serve to disclose the structure of the individuality. We can determine whether a performance of greater or lesser degree depends on talent or on intelligence; we can investigate what degree of correspondence exists between the experimental results and the teacher's judgments of the intelligence of pupils; we can delimit the extent to which general school efficiency is dependent on intelligence itself on the one hand and on non-intellectual factors on the other hand—a delimitation that forms one of the chief merits of the psychological methods.

The present writer's judgment on matters psychological is not to be considered. We shall let the psychologists speak for us here. Yet, careful study extending over a considerable time has brought the conviction that two notable, important and fruitful workers, my colleague Professor E. L. Thorndike and C. Spearman, of the University of London,. both, but especially the former, through their enthusiasm, exuberance, genius, and a rather natural carelessness of details and disregard of foundations and fundamental assumptions, have perpetrated on workers in other fields ideas, opinions, surmises, as scientifically established facts, in a way that a more scientific and careful worker could never have done. Only one of rare genius can touch with acceptable suggestions so many different fields of scholarly endeavor as has Professor Thorndike. Yet this very element of genius may have robbed his writings and public addresses of that desirable scientific flavor, and made it possible for psychologists of such note as Titchener, *l. c.*, p. 99, E. B. Holt, *The Jour. of Abnormal Psychology,* Jan. '16, pp. 366–7, and C. H. Judd, "Psychology of High School Subjects," to furnish us with such reviews, lamenting a general carelessness of exposition, obscurity, and even incoherence and equivocal expression. The writer would call attention to such implied and insinuated meanings as illustrated by the title of a paper presented at the Cincinnati meeting of the N. E. A., "Can algebra and geometry be reorganized so as to justify their retention for high school pupils not likely to enter technical schools?" and also by several sentences in the address on "Educational Diagnosis," *Science,* Jan. 24, 1913, pp. 133 sq.

Be it clearly understood that all this is not to be taken as an attack but as an analysis of the causes of the rather remarkable influence of the work of Professor Thorndike and those who have followed him. We all, the writer takes it, are working towards the same end, and with the same degree of sincerity, and therefore find neither pleasure nor interest in attempts at weakening another's position. We are all searchers for truth, regardless of what that truth reveals, and without preconceived ideas that we wish to establish. Relative to the statistician, let me say, that no scientific worker needs more to have the qualities of the ideal scientist, sincerity, honesty, open-mindedness, and zeal for truth. This is true because there is for him far greater temptation to try to corroborate preconceptions, to establish pet hypotheses.

A fellow-student with me at Johns Hopkins University was several years ago asked to call on an official of a large corporation that needed an expert statistician. After a bit of conversation the official said, as if in great surprise: "Did you not say that you are a statistician? But you are a scholarly gentleman." This official's conception of a statistician is alas all too common and it is strengthened when men who are scholarly and stand high in some line of research use the methods of statistics blindly or without understanding fully the fundamental principles and assumptions. The work of statistician should be raised from its low estate, and might be if all statisticians had to be certified as are accountants.

Reverting to the series of quotations, we shall consider the third, from pp. 11, 12: (*a*) It seems quite clear that those under discussion do not at least keep in mind the fact that the work is still in the preliminary stage, or rather did not when that was more the case. Much has had all the marks of being regarded as a diagnostic canon, and its influence has been all the same as if it was. (*b*) On this point the writer would refer to the Report of the Arithmetic Committee in this journal of March, 1915, pp. 125–6. The greatest amount of *joint* discussion on that report was had during the hour that three of the committee searched for a bite of lunch along Broadway just before the report was made. This furnished poor opportunity for me to secure more than modifications, so that "suggestions and recommendations," Nos. 1, 2 and 6 remained, under my protest for the reasons expressed in this "warning (*b*)," then unknown. The writer accordingly endorses most heartily the quotations from Stern, pp. 11, 12, 34, 40, and deplores the time worse than wasted by having every young school teacher conducting tests on the strength of a few weeks spent in a course in educational psychology or without any special training whatsoever. (*c*) Psychological tests have been overestimated, and have sometimes replaced and interfered with the teacher's getting the spirit of her subject and of transmitting it to her class.

Further on p. 34, we note that long before psychological tests were made, Pope observed what is a great but much neglected factor in the conduct of all such tests:

> The difference is as great between
> The optics seeing as the objects seen.
> All manners take a tincture from our own,
> Or come discolored through our passions shown.

This interaction between the tester and the tested as well as the fact that every one realizes more or less (and in an indeterminate degree) that he is being tested, make any pretense at precision in these so-called measurements futile. And if there is no claim to precision, why strain over the instruments of precision, the various coefficients, etc.? A line may be run and surveying done in the wilds of the Adirondacks with a Jacob's staff, but along Wall street the finest transit is required. The legitimate use of the instruments of precision is pointed out by Stern in quotations from pp. 8 and 26, the first two herein given.

A graduate student recently, in defending "mental measurements," said that they are now in the stage analogous to that of measuring the intensity of light when a candle was used as the unit. He did not recognize that the "international candle" is simply a far more refined unit, just as the standard meter is more refined than the length of a man's foot. Continuing his figure, we may add that in measuring light the matter is very much complicated when quality or color enters in. In psychological "measurements" the quality factor is preponderant but largely ignored. Spearman's plan of taking very few observations and making hypothetical general corrections on stated assumptions is a way that makes no appeal to a man who has the scientific method in his blood. To such jugglery, the latter quietly shakes his head and *sotto voce* says *"Ignoramus."*

The essential of measurement is nicely stated in Boole's "Finite Differences," pp. 106–7, thus: "In describing a magnitude there must be a standard magnitude of the same sort 'recognized alike by ourselves and those whom we address.' Of vague and indefinite expressions, 'very tall,' etc., we are never certain that those with whom we are conversing 'attach exactly the same idea to the expression that we do.' This vagueness is of little consequence in common life, because in most cases it is impossible to make an accurate estimate." This indispensable quality does not appear at all clearly in "mental measurements" nor yet in the formation of so-called *scales,* of which we shall speak briefly later. We can test a bar of steel in the laboratory, but we cannot measure the things of the spirit of man. One has said, "When I am weak, then am I strong." The testing goes on and the results seem regular, until the limit seems near, when lo, something else comes in. The person seems to get "second wind"; in weakness, comes a new flood of strength. This phenomenon is doubtless known to you all in personal experience. It has been put into exquisite verse by our "children's poet":

As torrents in summer,
Half-dried in their channels,
Suddenly rise, though the
Sky is still cloudless,
For rain has been falling
Far off at their fountains;

So hearts that are fainting
Grow full to o'erflowing,
And they that behold it
Marvel, and know not
That God at their fountains
Far off has been raining!

Nevertheless, *tests* have proven their use, some sorts without calculations of precision and others with the assistance of such forms of computation, but *all only when properly administered by scientifically trained experts.* The remaining quotations from Stern, pp. 49, 63, 64, and 9, were included to testify to this. Prof. S. A. Courtis has written, in *The Elementary School* Teacher, April, 1914: "In the writer's thinking, the functions of standard educational tests are four: (*a*) to secure information that will enable school authorities to formulate in objective terms the ends to be attained in any educational processes; (*b*) to measure the efficiency of methods designed to produce the desired results; (*c*) to determine the factors and laws which condition learning and teaching; (*d*) to furnish data that will enable comparisons of school with school, and teacher with teacher for purposes of supervisory control to be made upon *scientific, impersonal, objective* bases." The extent to which these desirable ends are being attained is not very clear. In his annual report, pp. 120–122, over date of Oct. 15, 1914, President Pritchett, writing on educational surveys, says: "There is, in fact, some danger that the faith in statistics may overshadow the main purpose of the school—the education of the pupils. The tendency to neglect the product by all-absorbing attention to the process is too marked at present to be overlooked." All three pages should be read to get his view of the state of these matters of supervision and administration.

We are brought thus naturally to give the promised examples of classes of errors and misconceptions found in the literature, for "Educational Administration—Quantitative Studies," Strayer and Thorndike, furnishes a good illustration of the way in which statistical data are used. On p. 100, a study of the social and economic status of teachers is begun. From it we quote, "Who are the teachers of our children? . . . What kind of teachers do we get for the money we pay? Is there any relation between the amount of salary a teacher receives and the amount of training secured by him? From what social group do teachers come? These and many other similar questions must be answered by anyone who would attempt to judge of the efficiency of our public systems of education. In the investigation by Professor L. D. Coffman, entitled 'The Social Composition of the Teaching Population,' we have the answer to our questions.

"Dr. Coffman's research is based upon the answers received to a questionnaire which was answered by 5,215 teachers selected at random in seventeen states. Most of the answers were secured from teachers who were in attendance upon their annual institutes. The purpose of the questionnaire was explained and replies were received from *all* of those present. . . . The tables are in the main self-explanatory.

"When we ask, Who are the teachers of our children, we must inquire concerning the families from which teachers come. The social status and the income of the parents of teachers limits the social inheritance which these teachers transmit to children. The following tables giving the occupations of parents, their income, and the number of brothers and sisters present a clear picture of the social and economic groups represented by the families from which teachers are recruited."

The affirmation that "we have the answer," and the statements, "selected at random" and "replies . . . from *all* . . . present," inserted to strengthen belief in the former, might be quite amusing if they were not intended to be serious, especially when we read (between the lines, in part) that most of the replies were from teachers assembled in their *county* institutes, from their rural schools in Middle Atlantic and Middle Western states. Rather "a clear picture" of this is obtained and corroborated when we see that 687 men out of 985 who stated the occupation of parent, and 1,409 out of 3,153 women who stated occupation, answered "farmers." What else would you expect? and what does it all mean? And then the assumption as to social status and income—well, what does every thoughtful student reader think as he studies the pages? We are sure Professor Thorndike would not for a moment think of claiming that the income of his father, a Methodist minister, and his own present annual income are in any way proportional to the real worth of the two men to their day and generation, or even of their social status in their respective communities. All real men do not know their worth and much less lay claim to it. A keen observation is portrayed in the parable in the asking, "When saw we thee sick, or in prison, and came unto thee?" You can't measure worth that way.

Examples might be multiplied to show that unwarranted, overdrawn or trivial conclusions are drawn, but we must close with but two more. Table 38, p. 131, relates to "Relative frequencies of different salaries in public and private secondary schools in the same localities. Percentages estimated from nineteen cities." Now a moment's reflection tells us that many of the country's best private secondary schools, paying the highest salaries, are not located in cities, and that the best schools, both public and private, are more likely not to be "in the same localities" for defective schools of one type fosters better schools of the other type. Of what real worth in drawing conclusions can such a table be? Surely we cannot speak of "accurate measurements" as is done elsewhere in the book. On p. 239 amongst other conclusions we find, "a large amount of time spent on arithmetic is no guarantee of a high degree of efficiency." Why must months of valuable time for pupils and teachers be spent to reach this profound conclusion, manifest to any thoughtful teacher or superintendent? It seems like a fifty-page sociological study that concluded that the Lower East Side is most densely populated and mostly by Jews, or like the sage conclusion that training in addition does not increase reasoning ability. Who with any power of introspection and observation ever claimed the latter untrue, and that without accusing himself of (p. 132) "the intellectual crime of giving mere opinions where indolence is the only excuse for failing to verify them." Surely nobody can charge the statistic users of indolence, yet the type of conclusions just cited recalls the excellent line of Horace ("Ars Poetica," l. 139) descriptive of some literary productions:

Parturient montes, nascetur ridiculus mus.

Turning next to "Theory of Mental and Social Measurements," E. L. Thorndike, The Science Press, 1904, we shall mention only the first two pages, saying that their statements and spirit are considered not expressive of their writer's present thought on the subject, that as they appear several statements are contrary to fact and their spirit is such as to misrepresent both mathematics and statistics to the reader. Professor Titchener (l. c., p. 99) has ably reviewed this work, and closes with the remark: "In any event, however, a mathematics for the psychologist has still to be written." Several years of teaching the subject has convinced the present writer that the sort of text most needed to fulfill the demand is one that he has been inspired, in part by the above challenge, to conceive and begin to write. In brief outline, it is to present a rigorous treatment beginning with the theory of the quadratic equation, leading to such topics in the general theory of equations as are directly useful in statistical work or in a rigorous development of other topics that are necessary; continuing the algebra through the topics that lead to a careful study of mathematical probability, it will treat that subject and its train of related-subjects, interspersing analytic geometry and the calculus before it is possible to complete that train. The second part is to treat as carefully the various statistical measures, constantly keeping clearly in view the fundamental assumptions made in their development. The book is to contain an exhaustive treatment of correlation that should clear the atmosphere at a point where blue sky has long been hoped for. Read the recent article by Prof. C. N. Moore in *School and Society,* Sept. 11, 1915. See also *Science,* Apr. 24, 1914, p. 609.

Illustrating further the kind of mathematical provender provided for the students of psychology, we cite "Experimental Psychology and Pedagogy," R. Schulze, translated by Rudolf Pintner, The Macmillan Co., 1912. Chapter I., on the Mathematical Treatment of Results, starts (p. 17) with the results of eighty measurements of the length of a certain rod. They are absolutely symmetrical, the frequencies being 1, 3, 8, 17, 22, 17, 8, 3, 1, the range being 0.4 mm. on 100 mm. The author must have practiced some time to have an actual set (so stated) of eighty observations come out so beautifully. Then on p. 19 follows: "Since in our measurements the accidental errors are symmetrically grouped, the arithmetical mean of all the single measurements will give us the real length of the measured object." Would we might ever know, and at least so easily, the *real* measure of any quantity! Further, in his case, the mean, median and mode coincide so there is no chance to state any comparative values of the three averages, as has been done by another but without proper understanding. From pp. 20, 21, we quote, "Therefore it is always best to calculate the average error," and "it has been established that the probable error always amounts to 4/5 (or, more precisely .8453) of the average error." "Always"—under what assumptions? Of these there seems to be utter disregard or ignorance. Finally on pp. 23 and 24 the crass superficiality and astonishing naïvete of the whole treatment appears in the paragraph:

"The formulæ that we have so far given are accurate enough when more than ten measurements are taken. Other formulæ have been found to be better for a smaller number of measurements. We give below the simple formulæ for reckoning the error that we have just discussed, and append thereto the more accurate formulæ (Nos. 1_a, 2_a, 3_a).

"THE SIMPLE FORMULÆ.

"1. Average error:

$$E_m = \frac{\Sigma\Delta}{n}.$$

"2. Probable error of the separate observations:

$$PE = \pm .8453\frac{\Sigma\Delta}{n}.$$

"3. Probable error of the arithmetical mean:

$$PE_m = \frac{PE}{\sqrt{n}}.$$

"MORE ACCURATE FORMULÆ.

"1_a. Average error:

$$E_m = \sqrt{\frac{\Sigma\Delta^2}{n-1}}.$$

"2_a. Probable error of the separate observations:

$$PE = \pm .6745\sqrt{\frac{\Sigma\Delta^2}{n-1}}.$$

"3_a. Probable error of the arithmetical mean:

$$PE_m = \frac{PE}{\sqrt{n}}.$$

"Note the fact that even with eighty measurements the accuracy of the simple formulæ is very great." The numerical values from the test are given in each case.

What rhyme or reason there is in dubbing the latter formumulæ "more accurate" is past our finding out. Their real meaning is known to all readers.

The meaningless indefiniteness of the following sentence will be apparent to all (p. 341): "If we define our problem clearly and carefully enough, if we do not make it too wide, the necessity of using that formula can be avoided." The carelessness of the author, or translator and others is shown on p. 347: "The only experiments on these lines known to the author are those carried out by Spearman in AMERICA."

Mention should be made of "Empirical Studies in the Theory of Measurement," E. L. Thorndike, The Science Press, April, 1907, mainly in the way of appreciation. Attention is called (p. 26) to the fact that the assumptions made in the formulation of the Pearson coefficient of correlation "never are except by chance more than approximately" fulfilled, "and in the majority of the cases in which students of the mental and social sciences need to measure relationships, they are far from true." As was stated above, we feel that wherein it is known that data do not fulfill the supposed conditions, it is vain to strive to uphold any particular measure. We should all realize the particular properties of each average and measure, and endeavor to present to our readers the maximum not the minimum of information as to the material under consideration. A timely book in this respect as to averages is the translation by Prof. W. M. Persons of "Statistical Averages" by Dr. Franz Žižek, Henry Holt and Co., 1913. On pp. 5, 6, Prof. Thorndike in comparing the average and the mean square deviation does so through the *particular* series of samplings he had been using, twice stating the fact. General not particular considerations furnish the basis of final judgment however.

The work of Dr. Spearman could profitably be considered far more than space and time permit. Brown, l. c., p. 89, in a footnote says: "It is perhaps hardly necessary to point out that the detailed consideration of Dr. Spearman's work given here and elsewhere in the present book is, and is meant to be, a tribute to the outstanding importance of that work. No one can appreciate more vividly than does the writer the valuable suggestiveness of all Dr. Spearman's writings, above all in respect to the special precautions therein emphasized as requisite in measuring *psychical* (as distinct from anatomical, etc.) correlation." He therein strikes the key-word, suggestiveness, which earlier led us to couple Professor Thorndike and Dr. Spearman. It seems to us however that the latter's work is more like Herbart's "constructions" to which we referred on p. 4, Vol. IX., No. 1, of this magazine. Any reader desiring to have a careful evaluation of Spearman's work should read Brown, pp. 54–133. In the Proceedings of the 6th International Congress of Psychology, Geneva, 1909, Brown presented a paper in which he attacked Spearman in part as follows:

"Thus the fallacy involved in Spearman's correlation formula is the assumption that variability of individual performance may be classed with errors of observation for the purposes of correlation—an assumption which does not conform with the facts. In view of the difficulties attaching to the correlation formula it is perhaps wiser to leave our original coefficients in their crude form, at least until a method capable of correcting for variability correlation has been devised. The remedy at present would seem to be to increase the number of original measurements of each mental capacity. . . . Much of the work done on correlation by psychologists in the past has been directed towards the determination of the closeness of relationship between elementary psychical capacities, the investigators using the variability of individual measurements as a stepping stone to the discovery of relations which they believed to be relatively constant. These relations, however, may themselves be found to be very variable. Correlation coefficients may—nay, do—vary from type to type, and these 'typical' coefficients with their relations to one another are likely to prove of greater importance in the science of psychology than those calculated from a mixture of many types. A great deal of work has yet to be done in the field of mental variation, in which correlation will be of continual service, before variations can be at all safely used as a guide to the constant causes or relationships underlying them. On the assumption that a correlation coefficient is a correct measure of the degree of the identity of causation of the two abilities correlated, the question of the numerical relations to one another of all the coefficients that may be calculated between all the possible pairs of abilities correlated is of considerable interest."

Spearman's reply may be found in *The (British) Journal of Psychology,* Vol. III., Part 3, Oct., 1910. His first two paragraphs call attention to "accidental" errors and claim their "size can be measured by the size of the discrepancies between successive measurements of the same thing." Either he does not give the technical meaning to "accidental" errors, or his statement does not agree with the findings in the physical realm. The decision seems to us of no great consequence for, applied to coarse, rough material as is the case in his field, the formula has no *clear* meaning at least. The contention impresses us as analogous to one between two men who, having *stepped* off twenty paces and with a cord staked off a circular plot, are now discussing whether 22/7 or 3.14159 should be used in computing the area. However, Spearman has rendered valuable service in calling to the attention of psychologists the need of caution and vigilance towards detecting hidden factors that effect the values of correlation coefficients. That same caution has likely made workers avoid to a greater extent the wrong idea that adjustment of observations by the method of least squares will give quality and tone to poor, indifferently made observations.

Finally, we bring to your attention a paper in *School and Society,* July 31, 1915, written under Prof. Hollingworth on "An Experimental Study of Self-Analysis, Estimates of Associates, and the Results of Tests." This particular paper is chosen as it is recent, illustrates a number of points, and keeps the matter in the family. At the end of the second column on p. 172 begins a paragraph that shows that greater care in expression that characterizes more and more some writers at least: "The results show only what happened in this case, and only to that degree do they suggest what we may expect to be generally true. Many similar studies must be made, under all sorts of conditions and by a variety of methods, before we shall have the final answers to our questions."

The method is stated of selecting 25 juniors in college, who at intervals of two to four weeks were given slips of paper bearing the names of all the group of mutually acquainted students, and requested to arrange them in an order of merit as to the possession of the traits, neatness, intelligence, humor, conceit, beauty, vulgarity, snobbishness, refinement and socialibity. A similar group of seniors judged themselves and one another in the same way. The groups were also given well-known psychological tests such as Completion, Association, Opposites, Mixed Relations.

In the second column, p. 173, we read "her median or true position" and in *The Jour. of Phil., Psy. and Scientific Methods*, Oct. 8, 1914, coming also from Barnard College (p. 578), "the average position for each item was calculated. The order of these average positions was considered the objective order," then "this objective or correct order." At the risk of being considered priggish, we call attention to these words "true" and "correct," because they illustrate a fundamental consideration, a considerable factor, in the influence that the psychological work has had. The insinuation of the false was the method of the Serpent of old; and yet, in more serious vein, we must confess that such insinuation, antipodal to the scientific method and spirit, always makes our spirit feel that liberties had been taken with it.

The above is closely followed by this:

"That is to say, the individual's error in judging herself is somewhat greater than the average error of her friends in their judgments of her. The individual does not judge herself *so* accurately as she is judged by her friends."

Surely faith in statistics has run away with both good judgment and common sense!

(P. 177.) "In the long run it is true that she who knows herself best is the best judge of others, or at least tends to be a good judge. The degree to which this is true, however, varies with the trait in question. With the 'undesirable' traits of snobbishness, conceit and vulgarity the coefficients are so low as to be quite unreliable and perhaps due only to chance. The same is true of neatness. But in the cases of refinement, humor, beauty, sociability and intelligence the coefficients are fairly high."

The twenty-five must arrange *all* for specified qualities—some desirable, some undesirable—when presumably some are deemed as not possessing certain of those qualities. How can I classify 24 colleagues say, as to their degree of guilt in the matter of beating their wives, when to the best of my knowledge none of them have ever thought of resorting to that as a form of recreation or of exercise of authority? The article as a whole took us back to a worn, dusty copy of *The Gettysburg Mercury,* our college monthly, wherein we had published some junior thoughts on "Import of Interpretations" of nature and of man's acts, noting that "There *is* a difference in the interpretations and that too dependent upon the former experiences of the observer. We readily recognize a Selective Absorption in mind as well as in matter." That "the subjective element is as manifest in the interpretations of man's acts as of nature," also that "some people see honor, integrity and uprightness wherever there is the slightest semblance of it. Why? Because it is what they would expect to find in their own acts under similar circumstances," and that "your true feelings, being known, enkindle the same feelings towards you," and finally "if you find yourself constantly adjudging men as honest and trustworthy or as false and hypocritical, you may be well assured that these are characteristic of your own life and that too to a greater degree than you may be ready to admit. Paul says, 'Wherein thou judgest another, thou condemnest thyself.' Our judgments have ever a tincture of our own character." David Hume in 1817 said: "There is a universal tendency among mankind to conceive all beings like themselves and to transfer to every object those qualities with which they are familiarly acquainted." Some such reflections, we feel, might have told as truly the results in the experiment as the statistical record without the expenditure of time and labor.

It gives us a sort of what's-the-use feeling to contemplate such experiments and when it comes to being told that our friends judge us more accurately than ourselves, we like to turn back again to Browning and refresh our spirit with

> Now who shall arbitrate?
> Ten men love what I hate,
> Shun what I follow, slight what I receive;
> Ten, who in ears and eyes
> Match me: we all surmise,
> They, this thing, and I, that: whom shall my soul believe?
>
> Not on the vulgar mass
> Called " work," must sentence pass,
> Things done, that took the eye and had the price;
> O'er which, from level stand,
> The low world laid its hand,
> Found straightway to its mind, could value in a trice:
>
> But all, the world's coarse thumb
> And finger failed to plumb,
> So passed in making up the main account:
> All instincts immature,
> All purposes unsure,
> That weighed not as his work, yet swelled the man's amount:
>
> Thoughts hardly to be packed
> Into a narrow act,
> Fancies that broke through language and escaped:
> All I could never be,
> All men ignored in me,
> This I was worth to God, whose wheel the pitcher shaped.

Starting with the assumption that some qualities are more indicative of "General Intelligence" than others and more comprehensive, we venture to suggest the value of tests that might get at such information as

1. The normal size of the angle of the cone of vision within which perception is clear.

2. The normal region of awareness in general and for specific classes of stimuli.

3. The power to reach decisions or "make up one's mind."

4. That given by Spencer, "Study of Sociology," p. 32—

"If you want roughly to estimate any one's mental caliber, you cannot do it better than by observing the ratio of generalities to personalities in his talk—how far simple truths about individuals are replaced by truths abstracted from numerous experiences of men and things."

The promised remarks on educational scales must hold over, and preface the study of correlation, coefficients and ratios, correction formulas, the methods of ranks and grades and the "Order of Merit Method."

SUMMARY.

The "first scientific attempt" by Mrs. Sophie Bryant, D.Sc. Lond., recognized matters of fundamental importance that, at least some, later workers would have done well to bear in mind.

The "movement against mathematics" that stirred up this long investigation, and the causes that influenced changes in curricula, were due in large part to the sanguine hopefulness, the exuberant suggestiveness of Prof. E. L. Thorndike and others who followed him, to the resultant impression and sometimes profession that certain matters had been scientifically established which in some cases are even less confidently affirmed to-day, and to the employment of the mathematical forms and methods of handling scientific data and the transfer to the work of the universal feeling of certitude or confidence in mathematics. Finally, this confidence was possibly strengthened for many by the refinements introduced by Dr. C. Spearman, of the University of London, the validity of which for most data to which they were in this field applied is, to say the least, yet to be proven.

Some search has not led us to disagree with Dr. Henry S. Pritchett in saying that there is "no science as yet of administration of educational affairs."

There have been tests that have in a very rough way given a first approximation to what may reasonably be expected in certain school grades in some of the school subjects. The "principle" of tests has possibly been vindicated as Stern has pointed out.

The great amount of time that has been spent on tests by those not scientifically qualified to administer them is deplorable. Qualified scientists employ enough time in arriving at the best without encouraging promiscuous testing by all sorts of novices.

The carelessness or indifference that permits stealthy implications in phraseology is also to be deplored.

Lastly, it is most earnestly recommended that all who have to teach mathematics make serious endeavor to get the spirit of the subject and a knowledge of those parts to which what they teach immediately leads so that thus you may be able to put to silence ill-founded attacks, and assist in spreading a better understanding of a science that can be made to induce to more general clear thinking on the every-day affairs of life.

Columbia University,
New York City.